72 ~~1~~. A method for encoding a sequence of video images, each video image comprising pixels, said method comprising the steps of:

representing the sequence by at least one set of consecutive images, each set of consecutive images comprising a plurality of video images;

representing each set of consecutive images by one reference image and a subspace model of motion, the reference image and the subspace model of motion together comprising an IDLE model, wherein the subspace model of motion comprises a plurality of temporal coefficient vectors and a plurality of basis images, so that each element of each temporal coefficient vector corresponds to one video image and so that each element of each basis image contributes to the motion for a pixel, the combination of one temporal coefficient vector and one basis image together being called a factor,

representing the sequence in encoded form by the collection of IDLE models,

wherein for each IDLE model the following steps are performed:

(1) selecting a first video image to be represented by the IDLE model,

(2) selecting a reference image, called I-image,

(3) for each image different from the reference image, called U-image, beginning with the first image, estimating motion between the I-image and the U-image, thereby yielding a motion field, until a last image to be represented by the IDLE model is reached,

(4) defining a set of consecutive images to be represented by the IDLE model as the plurality of images from the first image to the last image,

(5) computing a preliminary model of motion for the motion fields estimated for the set of consecutive images,

A1 cont...

(6) selecting a subset of factors from the preliminary model of motion building a subspace model of motion, and

wherein the subset of factors for the subspace model of motion is selected after each update of the subspace model and/or after the building of the subspace model is finished, and

wherein a factor is included in the subspace model if one or a combination of the following criteria is fulfilled:

A1 cont...

(1) a given number of factors is not exceeded,

(2) a number of factors depending on the number of images to be represented is not exceeded,

(3) a norm of its temporal coefficient vector is larger than a threshold,

(4) a norm of its basis image is larger than a threshold,

(5) a norm of the product of its temporal coefficient vector and its basis image, called the norm of the factor, is larger than a threshold,

(6) criteria (3) to (5) applied after compression,

(7) criterion (3) to (6) applied jointly on several subspace models with common thresholds,

(8) a norm of the factor is larger than the value at the knee point in the sequence of norm values of all factors of the subspace model, the norm values being sorted to achieve a monotonic order.

13 ~~2~~. The method according to claim 1, wherein the fidelity or quality criterion is satisfied, if the number of pixels or blocks of pixels predicted with bad fidelity or quality does not exceed a certain threshold.

14 ~~3~~. The method according to claim 2, wherein the quality is exploited by considering how noticeable prediction errors are, taking into account masking effects of the human visual system.

75 ~~4~~. The method according to claim 1, further comprising:

transmitting and/or storing the reference image and the subspace model of motion,

wherein the results transmitted and/or stored together comprise the representation of the sequence of video images, which can be used to calculate predictions by warping the pertaining reference image according to the pertaining motion fields rebuilt from the pertaining subspace model of motion.

76 ~~5~~. The method according to claim 1, wherein one or more of the I-image, the temporal weight coefficients and the basis images are compressed before being stored and/or transmitted, wherein an individual transformation is applied on each factor yielding comparable subspace model prediction errors for each factor.

A1 cont.

77 ~~6~~. The method according to claim 1, wherein the subspace model of motion is updated successively when each new motion field is estimated.

78 ~~7~~. The method according to claim 1, wherein said step of selecting a reference image comprises a method for adaptively selecting the reference image which comprises the following steps:

selecting the first image as reference image, and/or

selecting an image with a given distance from the first image as reference image, and/or

selecting an image as the reference image whose histogram has the highest similarity with an average or median histogram of a given range of images, and/or said method for adaptively selecting comprises following steps:

(1) setting the image following the first image as a current image,

(2) estimating motion between the first image and the current image, thereby producing a motion field,

(3) calculating a prediction for the current image by warping the first image according to the motion

field, or

calculating a prediction for the first image by warping the current image

(4) computing a fidelity criterion, a quality criterion, or a cost criterion for the prediction,

(5) repeating steps (2) to (4) for the following images by setting the current image to the next image as long as the computed criterion is satisfied or a maximal distance from the first image is reached,

(6) selecting the reference image as the last image for which the criterion was satisfied.

79 ~~8.~~ The method according to claim 1, wherein in step (3) the choice of the last image to be represented by the IDLE model is made dependent on one of or a combination of the following:

(1) a given distance from the first image or the reference image,

(2) minimum transmission costs,

(3) scene shifts,

(4) fidelity or quality of predictions according to either estimated motion or a subspace model of motion,

(5) how well a motion field for a new image fits to a subspace model of motion computed for earlier images,

(6) a given limit on the allowed number of images between two consecutive reference images.

A1 cont'

~~30~~ 80 ~~9.~~ The method according to claim 8, wherein step (4) comprises the steps:

(1) calculating a prediction for the U-image by warping the I-image according to the motion field, or calculating a prediction for the I-image by warping the U-image,

(2) computing a fidelity criterion, a quality criterion, or a cost criterion for the prediction,

(3) choosing the last image to be represented by the IDLE model as the last U-image satisfying the computed criterion.

81 ~~10~~. The method according to claim 1, wherein for each reference image a shape field is given, defining for each pixel whether the pixel should be part of the IDLE model.

82 ~~11~~. The method according to claim 10, wherein for each image to be predicted, the shape field is warped according to the subspace model of motion and the warped shape field is used to indicate the valid pixels in the prediction.

83 ~~12~~. The method according to claim 11, wherein for each image, an input shape field is given, indicating for each pixel of the image, whether the corresponding predicted pixel should be valid and this field is compressed and transmitted and/or stored using the warped shape field as a predictor.

A1 cont...

84 ~~13~~. The method according to claim 10, wherein one or more of the reference image or the subspace model of motion is compressed utilising the shape field to achieve higher compression ratios.

85 ~~14~~. The method according to claim 10, wherein a collection of a reference image, a corresponding subspace model of motion and a shape field together comprise a video object, and a prediction is made as a synthesis from two or more video objects.

86 ~~15~~. The method according to claim 1, wherein the IDLE model is spatially extended so that one or more predicted images represented by the IDLE model are fully covered when the reference image is warped according to the subspace model of motion.

37 ~~16.~~ The method according to claim 15, wherein the extending of the IDLE model comprises an extension of the reference image by warping pixels from the original images corresponding to uncovered pixels of the predicted images into the position of the reference image and extrapolating the subspace model of motion correspondingly if given in the reference position.

38 ~~17.~~ The method according to claim 16, wherein the extending of the IDLE model in the case of forward motion compensation for each image to be covered comprises the steps:

(1) calculating a motion field from the reference image to the image,

(2) predicting the image by forward warping the reference image according to the motion field, thereby producing a prediction,

(3) increasing the size of the reference image, thereby producing an enlarged reference image,

(4) increasing the size of the motion field to the size of the enlarged reference image, thereby producing an enlarged motion field,

(5) filling the enlarged motion field by extrapolation, thereby producing an extrapolated motion field,

(6) backward warping the part of the corresponding original image not covered by the prediction to the co-ordinate system of the enlarged reference image according to the extrapolated motion field,

(7) including the backward warped parts of step (6) as new parts of the enlarged reference image, thereby producing a new reference image, and

(8) increasing and extrapolating the basis images of the subspace model of motion correspondingly.

A1 cont...

89 ~~18~~. The method according to claim 16, wherein the extending of the IDLE model in the case of backward motion compensation for each image to be covered comprises the steps:

(1) calculating a motion field from the image to the reference image,

(2) increasing the size of the reference image, thereby producing an enlarged reference image,

(3) predicting the image by backward warping the reference image according to the motion field, thereby producing a prediction,

(4) forward warping the part of the corresponding original image not covered by the prediction to the co-ordinate system of the enlarged reference image according to the motion field, and

(5) including the forward warped parts of step (4) as new parts of the enlarged reference image, thereby producing a new reference image.

A1 cont.

90 ~~19~~. The method according to claim 17, wherein in step (1) the motion field is calculated from the subspace model of motion.

91 ~~20~~. The method according to claim 1, wherein the reference image of a first IDLE model is used to predict the reference image of a second IDLE model and the reconstruction of the second reference image is based on the prediction from the first reference image plus a residual which is compressed and transmitted and/or stored.

92 ~~21~~. The method according to claim 20, wherein the prediction is based on warping of the first reference image according to a motion field estimated between the first reference image and the second reference image and

wherein the motion field is compressed and transmitted and/or stored or

wherein the motion field is included in the subspace model of

motion of the first IDLE model.

93 ~~22~~. The method according to claim 20, wherein the residual is dampened by considering how noticeable prediction errors are, taking into account masking effects of the human visual system.

94 ~~23~~. The method according to claim 22, wherein the strength of the dampening is made dependent on one or a combination of the following:

(1) the temporal distance from the last I-image,

(2) the strength of the dampening of the previous P-image residual,

(3) the spatial locations of pixels or blocks of pixels within the residual.

95 ~~24~~. The method according to claim 94 ~~23~~, wherein in (3) the pixels or blocks of pixels may be selected randomly such, that only a certain percentage of pixels or blocks of pixels is dampened.

A1 cont.

96 ~~25~~. The method according to claim 1, wherein a plurality of images, called B-images, is represented by a first and a second IDLE model, whose corresponding sets of consecutive images overlap.

97 ~~26~~. The method according to claim 25, wherein the overlap of the sets of consecutive images of the two IDLE models consists of those images which do not have a sufficient fidelity or quality when predicted from a single IDLE model only.

98 ~~27~~. The method according to claim 25, wherein for each B-image, a blend field indicates for each pixel or block of pixels the contribution of the predictions from the first and the second IDLE model.

99 ~~28~~. The method according to claim 27, wherein the blend field is a bi-level field, one level indicating that the prediction from the first IDLE model is used, the other level indicating that the prediction of the second IDLE model is used.

100 ~~29~~. The method according to claim 28, wherein the bi-level blend field is compressed using a bi-level compression tool.

101 ~~30~~. The method according to claim 28, wherein the blend field is processed by one or a combination of the following:

(1) median filtering,

(2) replacing every value in the blend field, with a new value which minimises a cost function,

(3) replacing every value in the blend field, with a new value which minimises a cost function which is given by a weighted sum of the prediction fidelity or quality and the corresponding roughness values of the blend field,

(4) dithering.

A1 Cont.

102 ~~31~~. The method according to claim 27, wherein for each pixel or block of pixels of the B-image the blend field value is a real number λ between zero and one, defining each pixel or block of pixels of the B-image as a convex combination of the corresponding pixels or blocks of pixels of the predictions from the first and the second IDLE model, where the first is weighted by $1-\lambda$ and the second is weighted by λ.

103 ~~32~~. The method according to claim 31, wherein the blend field is calculated so that it minimises a cost function.

104 ~~33~~. The method according to claim 32, wherein the cost function is a weighted sum of the prediction fidelity or quality of the resulting B-image and the roughness of the blend field.

105 ~~34~~. The method according claim 31, wherein a subspace model is build for the blend fields corresponding to the B-images within the overlap of the sets of consecutive images of the two IDLE models, the subspace model of blend fields comprising a plurality of temporal coefficient vectors and a plurality of basis images, so that each element of each temporal coefficient vector

corresponds to one blend field, so that each element of each basis image contributes to the blend field for a pixel, the combination of one temporal coefficient vector and one basis image together being called a factor.

106 ~~35.~~ The method according to claim 34, wherein the temporal weight coefficients and the basis images of the subspace model of blend fields are compressed, and/or wherein the basis images are represented in low resolution, and/or wherein a residual is compressed and transmitted for the subspace model of blend fields, and/or wherein the subspace model of blend fields is updated successively for each B-image.

107 ~~36.~~ The method according to claim 34, wherein the subset of factors of the subspace model of blend fields is selected according to one or a combination of the criteria given in claim1.

A1 cont.

108 ~~37.~~ The method according to claim 34, wherein the blend fields are warped to a common position before the subspace model of blend fields is built.

109 ~~38.~~ The method according to claim 26, wherein for assessing the quality of the prediction it is taken into account how noticeable prediction errors are, taking into account masking effects of the human visual system.

110 ~~39.~~ The method according to claim 1, wherein a residual is computed as the difference between an original image and its prediction, and wherein the residual is compressed and transmitted and/or stored.

111 ~~40.~~ The method according to claim 39, wherein the residual is dampened by considering how noticeable prediction errors are, taking into account masking effects of the human visual system.

112 ~~41.~~ The method according to claim 1, wherein the IDLE model also comprises a subspace model of intensity changes, the subspace model of intensity changes comprising a plurality of temporal coefficient vectors and a plurality of basis images, so that each element of each temporal coefficient vector

corresponds to one video image and so that each element of each basis image contributes to the intensity change for a pixel, the combination of one temporal coefficient vector and one basis image together being called a factor.

113 ~~42.~~ The method according to claim 41, wherein for the prediction of each image in the set of consecutive images, the corresponding intensity change, rebuilt from the subspace model of intensity changes, is added to the reference image before warping according to the rebuilt motion field from the subspace model of motion,

or is added to the preliminary prediction produced by warping the reference image according to the rebuilt motion field from the subspace of motion.

A1 cont.

114 ~~43~~. The method according to claim 42, wherein the subspace model of intensity changes is computed by the following steps:

(1) for each current video image in the set of consecutive images performing steps (2) to (4),

(2) calculating a motion field between the reference image and the current image,

(3) predicting the reference image by warping the current image according to the motion field, thereby producing a prediction,

(4) subtracting the reference image from the prediction, thereby producing a difference image,

(5) computing a subspace model of intensity changes using the difference images produced in step (4).

115 ~~44.~~ The method according to claim 42, wherein the subspace model of intensity changes is computed by the following steps:

(1) for each current video image in the set of consecutive images performing steps (2) to (4),

(2) calculating a motion field between the reference image and the current image,

(3) predicting the current image by warping the reference image according to the motion field, thereby producing a prediction,

(4) subtracting the prediction from the current image, thereby producing a preliminary difference image,

(5) warping the preliminary difference image according to the motion field in the opposite direction as in step (2), thereby producing a difference image,

(6) computing a subspace model of intensity changes using the difference images produced in step (5).

116 ~~45~~. The method according to claim 42, wherein the subspace model of intensity changes is computed by the following steps:

A1 cont'

(1) for each current video image in the set of consecutive images performing steps (2) to (4),

(2) calculating a motion field between the reference image and the current image,

(3) predicting the current image by warping the reference image according to the motion field, thereby producing a prediction,

(4) subtracting the prediction from the current image, thereby producing a difference image,

(5) computing a subspace model of intensity changes using the difference images produced in step (4).

117 ~~46~~. The method according to claim 42, wherein the subspace model of intensity changes is computed by the following steps:

(1) for each current video image in the set of consecutive images performing steps (2) to (5),

(2) calculating a motion field between the reference image and the current image,

(3) predicting the reference image by warping the current image according to the motion field, thereby producing a prediction,

(4) subtracting the reference image from the prediction, thereby producing a preliminary difference image,

(5) warping the preliminary difference image according to the motion field in the opposite direction as in step (2), thereby producing a difference image,

(6) computing a subspace model of intensity changes using the difference images produced in step (5).

118 ~~47.~~ The method according to claim 43, wherein the preliminary difference images and/or the difference images are dampened by considering how noticeable prediction errors are, taking into account masking effects of the human visual system.

A1 cont.

119 ~~48.~~ The method according to claim 41, wherein the subset of factors of the subspace model of intensity changes is selected according to one or a combination of the criteria given in claim 1.

120 ~~49.~~ The method according to claim 1, wherein a first set of basis images, pertaining to a subspace model of a first IDLE model, is used to form a prediction for a second set of basis images, pertaining to a subspace model of a second IDLE model, the second set of basis images being represented as the sum of the prediction and a prediction error.

121 ~~50.~~ The method according to claim 49, wherein the prediction is represented by means of a transformation matrix, this matrix acting as a linear transformation between the first and the second set of basis images, whereby said transformation matrix results from the minimisation of a cost function.

122 ~~51.~~ The method according to claim 49, wherein the prediction is determined as a warped version of a linear transformation of the first set of basis images, the linear transformation being represented by a transformation

matrix, whereby said transformation matrix results from the minimisation of a cost function.

123 ~~52~~ The method according to claim 51, wherein warped versions of the first and/or the second set of basis images are used instead of the original sets of basis images.

124 ~~53~~ The method according to claim 49, wherein for each basis image in the second set a decision is made to use either the original basis image or the prediction and the prediction error, whereby the decision depends on one or a combination of the following criteria:

(1) fidelity or quality of the prediction,

(2) the amount of data required for the combined representation of prediction and prediction error in comparison to the original basis image.

A1 cont...

125 ~~54~~ An apparatus for encoding a sequence of video images, each video image comprising pixels, said apparatus comprising:

means for representing the sequence by at least one set of consecutive images, each set of consecutive images comprising a plurality of video images;

means for representing each set of consecutive images by one reference image and a subspace model of motion, the reference image and the subspace model of motion together comprising an IDLE model, wherein the subspace model of motion comprises a plurality of temporal coefficient vectors and a plurality of basis images, so that each element of each temporal coefficient vector corresponds to one video image and so that each element of each basis image contributes to the motion for a pixel, the combination of one temporal coefficient vector and one basis image together being called a factor,

means for representing the sequence in encoded form by the collection of IDLE models,

wherein for each IDLE model the following is provided:

(1) means for selecting a first video image to be represented by the IDLE model,

(2) means for selecting a reference image, called I-image,

(3) for each image different from the reference image, called U-image, beginning with the first image, means for estimating motion between the I-image and the U-image, thereby yielding a motion field, until a last image to be represented by the IDLE model is reached,

(4) means for defining a set of consecutive images to be represented by the IDLE model as the plurality of images from the first image to the last image,

(5) means for computing a preliminary model of motion for the motion fields estimated for the set of consecutive images,

(6) means for selecting a subset of factors from the preliminary model of motion building a subspace model of motion, and

A1 Cont.

wherein the subset of factors for the subspace model of motion is selected after each update of the subspace model and/or after the building of the subspace model is finished, and

wherein a factor is included in the subspace model if one or a combination of the following criteria is fulfilled:

(1) a given number of factors is not exceeded,

(2) a number of factors depending on the number of images to be represented is not exceeded,

(3) a norm of its temporal coefficient vector is larger than a threshold,

(4) a norm of its basis image is larger than a threshold,